THESIS

THE CONSTRUCTION AND EVALUATION OF A SURVEY GROUP TEST IN ELEMENTARY SIGHT-SINGING

John P. Breen

Master of Music Education

1949

M. Mus. Ed.
1949
Breen

BOSTON UNIVERSITY

COLLEGE OF MUSIC

Thesis

THE CONSTRUCTION AND EVALUATION OF A SURVEY GROUP TEST IN ELEMENTARY SIGHT-SINGING

by

John P. Breen

(Mus. B., New England Conservatory of Music, 1942)

submitted in partial fulfillment of the
requirements for the degree of
Master of Music Education

1 9 4 9

Boston University
College of Music
Library

Approved by

First Reader Kenneth G. Kelley
Professor of Music Education

Second Reader Keith D. Snyder
Asst. Professor of Music Education

TABLE OF CONTENTS

CHAPTER		PAGE
I	PURPOSE OF THE STUDY.....	1
II	REVIEW OF RELATED RESEARCH.....	3
III	CONSTRUCTION OF THE TEST.....	14
IV	ANALYSIS OF DATA.....	20
V	VALIDITY.....	33
VI	SUMMARY AND CONCLUSION.....	40
VII	SUGGESTIONS FOR FURTHER RESEARCH.....	42
BIBLIOGRAPHY.....		43
APPENDIX.....		46
	A. GROUP TEST ADMINISTRATION.....	46
	B. GROUP TEST SCORING.....	50
	C. INDIVIDUAL TEST.....	51
	D. ABSTRACT.....	52

LIST OF TABLES

TABLE		PAGE
I.	RANGE AND DISTRIBUTION OF TOTAL SCORES IN RELATION TO POSSIBLE SCORES.....	21
II.	RANGE OF SCORES -- SUB-TEST ONE.....	23
III.	RANGE OF SCORES -- SUB-TEST TWO.....	24
IV.	RANGE OF SCORES -- SUB-TEST THREE.....	26
V.	RANGE OF SCORES -- SUB-TEST FOUR.....	27
VI.	RANGE OF SCORES -- SUB-TEST FIVE.....	28
VII.	RANGE OF SCORES -- SUB-TEST SIX.....	29
VIII.	RANGE OF SCORES -- SUB-TEST SEVEN.....	30
IX.	RANGE OF SCORES -- SUB-TEST EIGHT.....	32
X.	SCORES ATTAINED BY SAMPLE ON GROUP AND INDIVIDUAL TESTS.....	35
XI.	SCORE DISTRIBUTION OF CRITERION GROUPS.....	38

CHAPTER I

PURPOSE OF THE STUDY

CHAPTER I

PURPOSE OF THE STUDY

The primary purpose of this study is the construction of an objective group test designed to measure several skills which form a functional part of sight-singing ability. In order to evaluate this test as an instrument for measuring sight-singing ability, a correlation will be made between scores obtained on this group test and scores on an individual test of sight-singing.

Both tests will consist of elementary sight-singing material, that is, at the level of difficulty of simple hymn tunes. Generally, the tonal and rhythmic patterns included are taught in the fourth and fifth grades, and reviewed in the sixth grade either before or coincidental with new material. The public schools of Needham, Massachusetts, the location of this experiment, complete the material included in the test by the middle of the fifth grade. During the sixth grade this material is frequently reviewed in the sight-singing of simple tunes.

A secondary purpose of this study is the discovery, through use of this test, of the elementary sight-singing proficiency of a group of students who have completed the first six grades of a public school music program.

Because this study was initiated at the beginning of the school year, it was decided that the seventh grade students would be more representative than the sixth at this time of year for the purpose mentioned above. Therefore, for the experiment, all one hundred and eighty seventh grade students from Needham Junior High School were tested.

To secure correlation criteria an individual sight-singing test was constructed and administered to a representative sample of the seventh grade population. The scores on this individual test form a criterion of the sight-singing ability of the group which constituted the sample. By making a correlation between scores on the two tests, we can determine the efficiency of the group test as a measure of sight-singing ability.

The justification for this study lies primarily in the fact that there is no standardized group test now in existence which purports to measure sight-singing with any significant degree of validity. One of the expressed objectives of our puplic school music program is the development of sight-singing ability. Yet we have no reliable, scientific instrument with which we can ascertain our degree of success.

The purpose of this study is to construct a group administered measuring instrument which can serve as a valid measure of sight-singing ability.

CHAPTER II

REVIEW OF RELATED RESEARCH

CHAPTER II

REVIEW OF RELATED RESEARCH

After extensive reading of expert opinion in the field of music education, it was found that much controversy existed concerning the value of standardized music tests. This controversy centered on the question of what these tests actually measured. It was frequently said that they measure the formal techniques and not the more important aspects of reading and appreciation.

William Larson, Chairman of the Music Education Department at Eastman School of Music, writing a review of the Providence Inventory Test in Music for "The 1940 Mental Measurements Yearbook" says the following:

> "In this test, as with most tests of this kind, it is questionable that an indication of proficiency in this type of factual learning necessarily means that the student will have the ability to read music, for capacities of musical talent, to which the more mechanical aspects of music reading do not directly relate are undoubtedly of prime importance and basic for musical response. The use of a test of this kind for comparative purposes before and after a period of training such as indicated above may have the

effect of directing the attention of the teachers toward the more mechanical aspects of music reading at the expense of the more musical values, thus emphasizing a procedure which is now entirely too prevelant in school music teaching in the opinion of many prominent music educators. This issue is of paramount importance in the use of this kind of test from the point of view of considering its validity as a measure of musical achievement."

Karl Gehrkens[1] when writing about tests says:

"The musician often scoffs at tests, just as he scoffs at grades, records, and efficiency. It is true that these things are frequently exalted by educators to a point beyond their deserts, and the musician, being usually more intuitive and therefore frequently seeing more clearly into the very heart of life than the average man, comes to feel that these mechanical phases of education are receiving more emphasis in proportion to their importance than the artistic or spiritual elements. In other words, grades, tests, and efficiency often come to have a larger place in school work than inspiration, guidance, and intellectual and spiritual nurture."

1. Karl W. Gehrkens, Music in the Grade Schools, pp. 218-219

Glenn Gildersleeve,[1] the author of the standardized music achievement test bearing his name, and the state director of music education in Delaware says:

"Some music teachers avoid tests altogether, believing that the newer education should emphasize happiness in living and that tests emphasize the disagreeable. Tests and examinations are not designed for punishment or for instilling fear of failure. Instead, they are to be used to help us teach a mastery of the things pertinent to better musical performance and fuller musical enjoyment.

There are three serious limitations of teachers in the use of tests.

The first limitation is the one just mentioned - the wrong conception of the purposes of testing held by too many music teachers. Understanding the real value of tests as individualized classroom drills will do much toward removing this limitation.

A second limitation lies in the inability of many music teachers to interpret the results of a test they have given and to show these results to a class in a simple graphical way. Taking an X-ray picture of a

1. Glenn Gildersleeve, "Standards and the Evaluation and Measurement of Achievement in Music," Thirty-fifth Yearbook of the National Society for the Study of Education, pp. 206

person's teeth is a splendidly worked out scientific technique, but the dentist must be able to interpret the picture and explain to the patient the condition found or the X-ray has no practical value. The science of testing is of little value to the teacher who does not understand it.

A third and serious limitation is found in the incompetence of music teachers in constructing homemade tests and drills. Tests and drills should be musical and interesting, they should function with each pupil, and they should emphasize only those things that are essential."

In their chapter on music tests, Brooks and Brown[1] make the following conclusions which are pertinent to this study:

1. "There is no test which measures comprehensively and entirely satisfactorily any one of the larger processes or abilities which represent the main curriculum objectives of modern music education, such as, for example, song singing ability, or ability to read music, or appreciation of good music.
2. There is need for an elementary school test of

1. Marion B. Brooks and Harry A. Brown, Music Education in the Elementary School, pp. 187-188

musical information.

3. A need exists for a test of ability to read music, but perhaps such a test cannot be made."

In order to get a first hand knowledge of available standardized achievement tests in music the following tests were examined.

Hillbrand Sight-Singing Test

This is a test which measures individual sight-singing performance. The test consists of six songs which each student must sing at sight. The examiner records errors on a folder containing the songs. The range of difficulty of the items is small and the scoring is subject to the ability of the examiner to detect flatting or sharping in pitch. It is obviously extremely difficult to develop reliable norms on this type of subjective scoring.

This test could conceivably be useful to an individual teacher without making use of the norms, but a similar home-made test could very easily be constructed.

Kwalwasser-Ruch Test of Musical Accomplishment

This test is in ten parts: 1, knowledge of musical symbols and terms; 2, recognition of syllable names; 3, detection of pitch errors in a familiar melody; 4, detection of time errors in a familiar melody; 5, recognition of pitch names; 6, knowledge of time signatures; 7, knowledge of key

signatures; 8, knowledge of note values; 9, knowledge of rest values; and 10, recognition of familiar melodies from notation. This test is of the diagnostic inventory type. It is designed for use in grades four to twelve inclusive. Norms are given in deciles for each grade from four through eight and a composite norm for grades nine through twelve. The validity of this test rests primarily on the specifications for a standard course in music for graded schools adopted by the Music Supervisors' National Conference in 1921. The reliability, through chance-half technique, was found to be .97.

Knuth Achievement Tests in Music

This test purports to give "a valid and reliable measure of the pupils ability to recognize and comprehend music from its notation." The test has three divisions each designed for a different age group. Each division consists of forty items of the four-response multiple choice type.

The examiner plays a four measure phrase of which the first two measures and four choices for the final two measures are given. The student must pick the correct final two measures according to the rendition.

Validity is based on a textbook analysis of nine

public school music series, and expert judgment of six music supervisors. Reliabilities for each division were computed by using the Spearman-Brown formula. Norms in percentile are given for each division and for high school students according to musical background.

Providence Inventory Test in Music

This test is in ten parts: 1, naming notes; 2, placing "do"; 3, naming note values; 4, naming key signatures; 5, naming measure signatures; 6, naming rest values; 7, naming syllables; 8, naming melodies; 9, naming syllables (Bass clef); 10, naming symbols.

This test is designed for grades four through nine. The total time for administration is twenty-eight minutes. Norms are available in deciles for each grade.

Gildersleeve Musical Achievement Test

This test is in five parts: 1, recognition of melodies played by the examiner; 2, detecting changes in pitch, meter, key signatures; 3, writing key signatures, locating "la" in six keys, use of accidentals, knowledge of note values, knowledge of measure signatures, transposing a phrase from treble to bass clef; 4, fifteen multiple response questions on musical theory and history; 5, recognition of melodies from notation.

This is a survey type test measuring musical achievement

within the limits already stated. Tentative medians on three thousand cases are available.

The research work in the field of music education is limited, but the studies mentioned in the following paragraphs were found helpful as a guide in the test construction.

Raymond M. Mosher[1] in an attempt to find which battery of group music achievement tests would correlate the highest with an individual sight-singing test constructed and administered seven group tests to seven hundred and twenty-three children. The resulting correlations showed that the seven tests in the order of their power to measure sight-singing ability were:

1. Dictation of tonal patterns and figures.
2. Silent reading of melodies.
3. Dictation of melodies.
4. Dictation of rhythmic patterns and figures.
5. Recognition of scales and chords.
6. Knowledge of measure and note values.
7. Knowledge of musical symbols and terms.

A multiple correlation coefficient of .68 was obtained between group tests and the individual sight-singing test.

1. Raymond M. Mosher, A Study of the Group Method of Measurement of Sight-Singing.

Charles Dean[1] in determining the value of the Seashore tests for predicting success or failure in the required courses in sight-singing and ear-training at the Eastern Montana State Normal School obtained these results from two consecutive entering classes. With the teacher's marks as criteria, intelligence scores were found of little value in predicting success in sight-singing. Prior music training was of slightly greater value. The Seashore tests of pitch and tonal memory were the best individual predictors. These correlated with the criteria (music marks) to the extent of $.61 \pm .03$ and $.63 \pm .03$ for the first year, and $.62 \pm .03$ and $.70 \pm .02$, respectively, for the second year. The correlation results obtained between the criteria and the Seashore tests of consonance, time, rhythm, and intensity were all too low to be of material value in predicting sight-singing success. By combining the Seashore pitch, tonal memory, and years of previous musical training scores, and correlating the result with the criteria, a coefficient of $.77 \pm .024$ was obtained.

In a study[2] designed to find some reliable predictor of sight-singing success, which was made with one-hundred and thirty-one entering students at Washington State Normal School,

1. Charles D. Dean, "Predicting Sight-Singing Ability in Teacher Education", Journal of Educational Psychology, November 1937.

2. Frank Salisbury and Harold B. Smith, "Prognosis of Sight Singing Ability of Normal School Students", Journal of Applied Psychology, October 1929.

the following results were obtained. Single correlations between scores on an individual sight-singing test, as criteria of sight-singing ability, and scores on the Seashore tonal memory and pitch tests, gave coefficients of .641 and .604, respectively. A correlation between the same criteria and the score on a tonal dictation test resulted in a .596 coefficient. A year later the scoring on the individual sight-singing test was changed so that it became more objective and easier to administer. One-hundred and forty-four entering students were tested with these results. Single correlations between the sight-singing criteria and a tonal dictation test was .785; the criteria and Seashore tonal memory test - .656; the criteria and Seashore pitch test - .645. During both years correlations between the criteria and the remaining Seashore Tests scores were all much lower than those already mentioned and of no predictive value. Also, a measure of factual knowledge in music proved of little prognostic value.

From these studies we may conclude that the ability to remember tonal patterns, and the ability to distinguish pitch differences are both closely related to successful sight-singing. We may also conclude that success in tonal dictation is a good index of success in sight-singing. The research revealed no other studies directly related to the problem.

Through examination of several standardized music

achievement tests, it was found that none of them purported to be a measure of sight-singing ability. The opinion of many prominant music educators indicates that existing standardized tests are not sufficiently specific in what they measure. Also, that these tests measure very little music achievement other than factual knowledge concerning music. Some educators believe that conflicting pedagogical opinions in music education make the construction of a standardized music achievement test a difficult task. For this reason test constructors in standardizing a test for a large school population are apt to make too extensive use of factual material.

The problem then was that of constructing a group test which would give as nearly as possible the same results as an individual test of sight-singing.

CHAPTER III

CONSTRUCTION OF THE TEST

CHAPTER III

CONSTRUCTION OF THE TEST

An expressed purpose of the proposed group test is its use as an aid in discovering the number of students who are able to sight-sing simple hymn tune material. In order to define the limits of a simple hymn tune a study was made of three-hundred hymns from the English Hymnal. A frequency count of melodic intervals and rhythmic figures revealed that two-thirds (2/3) of the hymns studied did not exceed the following limits.

Melodic Intervals	Durative Values
1. All diatonic scale steps.	1. One tone sustained one beat.
2. All ascending and descending thirds between scale tones.	2. One tone sustained two beats.
3. Ascending fourths: 1-4, 2-5, 3-6, 5-8.	3. One tone sustained three beats.
4. Descending fourths: 8-5, 6-3, 3-7.	4. One tone sustained four beats.
5. Ascending fifth: 1-5.	5. Two tones to one beat.
6. Descending fifths: 8-4, 5-1, 2-5.	6. One tone sustained one-half beat.
7. Chromatic step 3-4#.	7. A dotted note to express a tone sustained for one and one-half beats.
	8. Rests equal to one-half, one, two and three beats.

For this study the hymn tunes which did not exceed the above limits shall be considered simple hymn tunes. The test, therefore, shall not exceed these limits.

A textbook analysis of those elementary music series in general use revealed that a knowledge of most of the melodic and rhythmic material mentioned above is taught in the fourth grade or before. By the middle of the fifth grade all the above material has been introduced.

Although at midyear it may be expected that all the fifth grade children know and understand this material, it is not expected that individually they will have developed skill in the use of this knowledge. Some of the children, usually the more musically gifted, will have developed a large measure of skill, others less, and still others almost none at all. Psychological studies show us that any skill or complex activity such as sight-singing is not developed through acquiring a knowledge of its individual parts. True, this knowledge must come first, but only through repeated absorption in the activity of sight-singing will a high level of skill develop. This test is designed to measure the amount of skill in sight-singing.

Sub-test One deals with recognition of note and rest values. It is set up as a completion test, the student being required to fill in the blanks with the value in number of beats. In order to do this, he must understand and correctly

interpret the measure sign.

In the light of the research this sub-test, which appears to be more a test of knowledge than of skill, would have a low correlation with sight-singing success, but it was inserted for diagnostic purposes. This may also be said of sub-test two and to a less degree of sub-test three.

Sub-test Two is a test of the ability to find the position of "do" from the key signature. The student is asked to place a cross on the correct line or space for each of six key signatures. This sub-test is based on the moveable "do" method of sight-singing which is used throughout the test. Obviously this would not be a valid test for a group who had not been taught to use moveable "do".

Sub-test Three tests the ability of the student to write the syllable names for four groups of notes, each group in a different key. There are some methods of teaching reading which do not use syllables. There are no statistics on this, but it seems safe to say that a large majority still use syllables. The children who were tested for this experimental study have been taught to use syllables.

Sub-test Four is designed to test the students auditory perception of twenty simple melodic intervals. Nineteen of the intervals used are diatonic and one is chromatic. The diatonic intervals are those which occur most frequently as revealed in the hymn study. The chromatic interval is

the third to the raised fourth (3/4#) in the major scale. It was found that this was the most frequently occurring chromatic interval.

This sub-test is in the form of a completion test. The syllable name of the first pitch to be played is given, and horizontally across from it a blank space is to be filled by the student with the correct syllable name of the second pitch. All the intervals are played in one tonality which is definitely established before the first interval is played.

In the research we discovered that tonal memory and the ability to distinguish pitch differences correlate highly with successful sight-singing. In this sub-test the student must perceive the pitch difference, keep the tonality in mind, and knowing the first give the syllable name for the second of the two pitches. In doing this the student is exercising tonal memory. He has been taught, in reading readiness and early reading, the pitch relationships of the major scale and has frequently sung these intervals. In developing sight-singing skill the teacher made use of this material in tonal dictation exercises.

Sub-test Five tests auditory visual perception of simple rhythmic patterns. It is set up as a four choice multiple response test with one example and eight items. The

student is required to circle the correct rhythmic pattern after hearing it sung.

Sub-test Six is closely related to the previous test. It makes use of the same rhythmic patterns and like the previous test is designed to measure auditory perception. Here the student is required not only to perceive the rhythmic pattern and recognize its notation, but also to recall how it appears in notation and write it. The ability to recall is generally considered a higher level mental activity than recognition. Although no studies are available, it appears that sight-singing activity requires more of the mental recall activity than the simpler recognition activity. The research does show that the Seashore tonal memory test results correlate highly with successful sight-singing. This may have some significance despite the fact that this is a specialized type of recall.

Sub-test Seven is a measure of the students silent reading ability. The test consists of ten musical motives. The student is required to read each motive silently and immediately following listen to this motive or one nearly the same played on the piano. The student must decide whether the motive played is the same or different from the one written.

Sub-test Eight is also a measure of silent reading

ability. The student is given the score of a complete hymn tune. This hymn tune is played only once with obvious mistakes in pitch and rhythm. While the hymn is played the student must follow the score and put a cross through any measure which is played incorrectly. The student is not allowed time to read through the music before it is played, but must read at sight simultaneously with the playing.

The test was given a preliminary tryout on thirty eighth grade students to determine ease of administration, clarity of directions, and to obtain student responses which might effect an improvement in the test. This procedure appeared advisable before proceeding with the experiment which involved a greater number of students.

The results of this preliminary tryout pointed out several areas in which changes were necessary. These changes were incorporated in the final form of the test as it appears in the appendix. In this form the test was given to the total seventh grade population of one hundred and eighty students.

CHAPTER IV

ANALYSIS OF DATA

CHAPTER IV

ANALYSIS OF DATA

The test described in the preceding chapter was built to survey the amount of development in sight-singing skill with elementary material. An analysis of the data from this point of view will center around the amount of accomplishment in the areas tested. However, if the test is to fit the need of a large population, it must be sufficiently difficult and at the same time sufficiently easy to serve the needs of varying stages of development in sight-singing skill. Therefore, the data is interpreted with this in view.

A table of the total scores in relation to possible scores gives a good overall picture of the group accomplishment. The highest possible score is one hundred and eighty-four (184). The highest score obtained by anyone in the group was one hundred and sixty-eight (168). The lowest score obtained was fifty-three (53).

TABLE I

RANGE AND DISTRIBUTION OF TOTAL SCORES IN RELATION TO POSSIBLE SCORES

Score Intervals	Frequencies
165-171	1
158-164	4
151-157	7
144-150	8
137-143	4
130-136	6
123-129	12
116-122	15
109-115	13
102-108	18
95-101	14
88- 94	24
81- 87	14
74- 80	15
67- 73	15
60- 66	8
53- 59	2
Mean	103.45
S. D.	26.6

NORMAL PERCENTILE CHART

By Arthur S. Otis

	Grade or group	No. of cases	Measure (Examination)	Form	Date	Examiner	Graphs by	School	City
Variable I	Sub-test IV								
Variable II	Total test scores								

Variable I Score intervals	Freq-'ncies	Sub-totals	Per cents	Variable II Score intervals	Freq-'ncies	Sub-totals	Per cents
38	5	180	100.				
36	7	175	97.2	165-171	1	180	100.
34	9	168	93.3	158-164	4	179	99.4
32	9	159	88.3	151-157	7	175	97.2
30	13	150	83.3	144-150	8	168	93.3
28	10	137	76.1	137-143	4	160	88.9
26	20	127	70.6	130-136	6	156	86.7
24	23	107	59.4	123-129	12	150	83.3
22	15	84	46.7	116-122	15	138	76.7
20	15	69	38.3	109-115	13	123	68.3
18	13	54	30.0	102-108	18	110	61.1
16	10	41	22.8	95-101	14	92	51.1
14	6	31	17.2	88-94	24	78	43.3
12	12	25	13.9	81-87	14	54	30.0
10	5	13	7.2	74-80	15	40	22.2
8	4	8	4.4	67-73	15	25	13.9
6	2	4	2.2	60-66	8	10	5.6
4	2	2	1.1	53-59	2	2	1.1
Median	24.			Median	101.0		

Standard Deviation Scale −3σ −2σ −σ M +σ +2σ +3σ

Published by World Book Company, Yonkers-on-Hudson, New York, and Chicago, Illinois. Copyright 1938 by World Book Company. Copyright in Great Britain. All rights reserved. PRINTED IN U.S.A. ONPC-7

For the scores on the group test to have meaning in terms of sight-singing skill the data in Table I should be studied in conjunction with the validation data in the next chapter.

Without considering the validity we can examine these scores purely from the standpoint of scientific test construction. Following this procedure we find that the distribution of scores very nearly forms a normal curve. This is graphically represented on the "Normal Percentile Chart". On this chart the normal curve is represented by a straight line.

The line representing the scores on this test is nearly straight between the tenth and ninetieth percentile points, which indicates a normal distribution for this group. In determining normality of a line representing only a few hundred cases, scores above the ninetieth and below the tenth percentile are not considered.

The normality of the curve representing the scores on this test indicates that the test was a good measuring instrument for this group.

TABLE II

RANGE OF SCORES -- SUB-TEST ONE

Score	Frequency
18	30
16	25
14	14
12	24
10	21
8	23
6	17
4	17
2	5
0	4
Median	12
Mean	11.1
S. D.	5.1

Inasmuch as this sub-test appeared to be a test of knowledge, which is generally considered a prerequisite to sight-singing, the mean score was not as high as was expected. The median score of 12, which represents three errors, is a better indication of the achievement of the total group. Fifty-five students or 30.5% of the group attained scores of 18 or 16, that is a perfect score or one error. The test offered no challenge for this group, but

further study shows a scatter of scores all the way down to zero. This probably indicates that the test fits the needs of the whole group.

TABLE III

RANGE OF SCORES -- SUB-TEST TWO

Score	Frequency
6	73
5	23
4	15
3	25
2	12
1	16
0	16
Median	5
Mean	4.0
S. D.	2.1

As expected, the median score of 5 on this sub-test was high.

Ninety-six students, or 53.3% of the group, attained perfect scores or made one error which in this situation was probably due more to carelessness than lack of knowledge. However, in the performance of a skill this carelessness is to be reckoned with.

Sixteen students or 8.9% of the group were unable to make a score. Another 8.9% of the group were able to get only one of the six items correct. This may indicate that these students who have been taught this simple bit of knowledge have forgotten it through lack of use, which may indicate disinterest because of little success in sight-singing.

TABLE IV

RANGE OF SCORES -- SUB-TEST THREE

Score	Frequency
20	28
19	2
18	1
17	1
16	4
15	24
14	10
13	4
12	1
11	8
10	14
9	4
8	3
7	3
6	2
5	12
4	8
3	7
2	4
1	11
0	29
Median	10
Mean	9.5
S. D.	7.1

Examination of Table IV reveals that the highest frequencies occurred at zero, 5, 10, 15 and 20. This is to be expected because the student who places "do" incorrectly will, if consistent, incorrectly name all five notes in that particular key. The scores which fall between these points are probably due to carelessness.

Seventy students or 38.8% attained a score of 14 or higher, which indicates few mistakes and relatively high achievement. Forty students or 22.2% attained a score of zero or 1. This latter group indicate very little knowledge or skill in the use of syllables as an aid to sight-singing.

TABLE V

RANGE OF SCORES -- SUB-TEST FOUR

Score	Frequency
40	0
38	5
36	7
34	9
32	9
30	13
28	10
26	20
24	23
22	15
20	15
18	13
16	10
14	6
12	12
10	5
8	4
6	2
4	2
2	0
0	0
Median	24
Mean	23
S. D.	7.8

An examination of Table V leads one to suspect a normal curve as resultant of the distribution of scores. The scores were plotted on a "Normal Percentile Chart" (page 22) and the resulting curve indicated a normal distribution. This plus the fact that no one obtained a perfect score indicates that this sub-test is a good measuring instrument for the whole group.

TABLE VI

RANGE OF SCORES -- SUB-TEST FIVE

Score	Frequency
16	8
14	22
12	40
10	42
8	40
6	20
4	6
2	0
0	2
Median	10
Mean	10
S. D.	3

TABLE VII

RANGE OF SCORES -- SUB-TEST SIX

Score	Frequency
16	0
14	1
12	10
10	17
8	19
6	19
4	40
2	43
0	31
Median	4
Mean	4.5
S. D.	3.6

Because the contents of Sub-test V and Sub-test VI are the same, Table VI and Table VII are studied together.

A study of Table VI reveals a normal spread of scores. One hundred and twenty-two students or 67.7% attained scores between 8 and 12 inclusive. Thirty students or 16.6% attained scores higher than 12, including 8 perfect scores. Thirty students or 16.6% attained scores lower than 12 including 2 zero scores. This sub-test, although limited in range,

apparently fits the need of the whole group.

A study of Table VII reveals low accomplishment in contrast to Table VI. One hundred and fourteen students or 63.3% attained scores of 4 or lower. Thirty-one students or 17.2% were scored zero. Because of the large percentage of low scores this sub-test, as a standardized test, was too difficult for the whole group, but it served us in indicating relatively low accomplishment in this area.

TABLE VIII

RANGE OF SCORES -- SUB-TEST SEVEN

Score	Frequency
20	1
18	2
16	10
14	28
12	37
10	36
8	38
6	19
4	8
2	1
0	0
Median	10
Mean	10.4
S. D.	3.3

Table VIII shows a good spread of scores which gives indication of a normal distribution. One hundred and eleven students or 61.7% of the total group attained scores between 8 and 12 inclusive. The scores on this sub-test indicate that for the most part it fits the need of the entire group. A few items which would give a sharper differentiation among the 61.7% mentioned above would improve the test.

TABLE IX

RANGE OF SCORES -- SUB-TEST EIGHT

Score	Frequency
48	2
44	13
40	18
36	23
32	44
28	33
24	25
20	19
16	3
12	0
8	0
4	0
0	0
Median	32
Mean	31
S. D.	7.2

An examination of Table IX reveals a fairly high median score. Thirty-three students or 18.3% of the group attained scores of 40 or higher. The distribution of scores appear fairly normal. The test is well within the ability of most of the group and fits the need of the entire group.

CHAPTER V

VALIDITY

CHAPTER V

VALIDITY

In the preceding chapter we discovered that the test discussed was a good measuring instrument for the group as a whole. The test was designed to measure several abilities which, according to the research, are indicators of skill in sight-singing. The efficiency, or validity, of the test as a measure of sight-singing is yet to be determined. That is the purpose of this chapter.

Several of the sub-tests were designed to measure specific abilities which have a significant correlation with sight-singing skill, as proved by the previous studies quoted in the research. This indicates some validity for the whole test, but by means of statistical techniques we can obtain a more accurate estimate of this validity.

A sample of the seventh grade population was chosesn. In order to insure this to be a truly representative sample of the tested population, all one hundred and eighty test booklets were arranged according to scores from the highest to the lowest and every fourth booklet taken for the sample. The sample thus obtained constituted 25% of the total seventh grade.

An individual sight-singing test was constructed within

the same limits of difficulty as the group test. (Included in appendix) This test was administered individually to the forty-five students who constituted the sample. In order to obtain the best possible objective score from each student in the sample, the measure was taken as a unit of score. That is, each measure sung correctly is scored two points, one for correct pitch and one for correct rhythm. This scoring technique was used successfully by Mosher.[1] It was also used during the second year of the experiment by Salisbury and Smith,[2] and is recommended by them as superior to other identical purpose scoring techniques which they tried with less success.

1. Raymond M. Mosher, A Study of the Group Method of Measurement of Sight-Singing.

2. Frank Salisbury and Harold B. Smith, "Prognosis of Sight-Singing Ability of Normal School Students", Journal of Applied Psychology, October 1929.

TABLE X

SCORES ATTAINED BY SAMPLE ON GROUP AND INDIVIDUAL TESTS

Student	Group Test	Individual Test	Student	Group Test	Individual Test
1	168	71	24	97	9
2	158	70	25	96	24
3	155	71	26	96	0
4	150	69	27	94	38
5	145	57	28	92	3
6	141	69	29	92	40
7	134	45	30	91	10
8	130	56	31	90	17
9	127	12	32	88	8
10	124	46	33	87	7
11	123	14	34	82	5
12	121	62	35	81	1
13	119	12	36	80	8
14	117	48	37	78	0
15	116	21	38	76	3
16	114	53	39	74	1
17	113	0	40	72	2
18	109	30	41	70	0
19	108	37	42	68	6
20	106	15	43	67	0
21	104	20	44	66	3
22	103	48	45	63	4
23	101	10			

The two sets of scores which appear in Table X were correlated and the product-moment "r" obtained. The formula used for the computation is as follows:

$$r_{xy} = \frac{\frac{\Sigma(fx'y')}{N} - c_x c_y}{\sqrt{\frac{\Sigma(fx'^2)}{N} - c_x^2}\ \sqrt{\frac{\Sigma(fy'^2)}{N} - c_y^2}}$$

$$r_{xy} = \frac{\frac{1202}{45} - 5.12}{\sqrt{\frac{1278}{45} - 5.52}\ \sqrt{\frac{1522}{45} - 4.75}}$$

$$r_{xy} = .838$$

An examination of Table X reveals an accumulation of twelve scores between zero and five on the individual test. These same students attained scores on the group test which spread between sixty-three and one hundred and thirteen. Five students scored zero on the individual test, and showed wide difference on the group test. This lack of spread in the low scores indicates the need of a more discriminating individual measure for the low scoring area if this method

of correlation is to yield the best results. The product-moment method of correlation is based upon the assumption that such relationships as exist between variables can be adequately described by a linear or straight-line relationship. That is, for each successive change or difference in the measures of one variable, there is a proportionate change in the other variable. We assumed when giving the individual test that variation in sight-singing skill could be expressed in terms of a linear scale or a continuum. The individual test proved that such an assumption needs further investigation.

The scores on the individual test distribute in a pattern which can easily be divided into three groups. We shall identify these groups from highest to lowest as skilled, some skill, and little or no skill. The thirteen students who scored forty-five and above will be placed in the highest group. Fifteen students who scored between nine and forty, inclusive, will be considered in the middle group and the remaining seventeen students who scored between zero and eight will be placed in the lowest group. We can now correlate the group test scores with the group ratings by means of a triserial correlation. This triserial correlation is based on the assumption that the data of the segmented variable are derived from a variable that is normally distributed.

TABLE XI

SCORE DISTRIBUTION OF CRITERION GROUPS

Test Scores	Criterion Ratings			Total Group
	No Skill	Some Skill	Skilled	
	f_L	f_c	f_h	f_t
165-171			1	1
158-164			1	1
151-157			1	1
144-150			2	2
137-143			1	1
130-136			2	2
123-129		2	1	3
116-122		2	2	4
109-115	1	1	1	3
102-108		3	1	4
95-101	1	3	0	4
88- 94	2	4		6
81- 87	3			3
74- 80	4			4
67- 73	4			4
60- 66	2			2
	N = 17	15	13	45

A study of Table XI revealed the following values:

M_L	= 79.9	y_L	= 0
M_c	= 104.5	y_c	= .381
M_h	= 135.7	y_h	= .342
p_L	= .38	M_t	= 104.4
p_c	= .33	σ_t	= 26.6
p_h	= .29		

The formula for triserial "r" is as follows:

$$r_{tri.} = \frac{y_h M_h + (y_c - y_h) M_c - y_c M_L}{\sigma_t \left[\frac{y_h^2}{p_h} + \frac{(y_c - y_h)^2}{p_c} + \frac{y_c^2}{p_L} \right]}$$

$$= \frac{.342(135.7) + (.381 - .342)104.5 - .381(79.9)}{26.6 \left[\frac{(.342)^2}{.29} + \frac{(.381 - .342)^2}{.33} + \frac{(.381)^2}{.38} \right]}$$

$$= \frac{20.05}{21.01}$$

$$= +.954$$

This correlation coefficient of +.954 is significant as an indication of the validity of the group test as a measure of sight-singing.

CHAPTER VI

SUMMARY AND CONCLUSION

CHAPTER VI

SUMMARY AND CONCLUSION

The purpose of this study was the construction and evaluation of a group test designed to measure several abilities which form a functional part of sight-singing skill.

A secondary purpose was the discovery of the sight-singing proficiency of the seventh grade pupils who made up the experimental group.

The group test was administered to the seventh grade population of Needham High School, Needham, Massachusetts. The data obtained from this administration was analyzed statistically with the following conclusions:

1. The test proved to be a good measuring instrument for the group.
2. The highest achievement was in the first, second and third sub-tests, which were primarily tests of knowledge.
3. The lowest achievement was in sub-test six which appeared too difficult for the group as a whole. Because of a large number of low scores, this sub-test did not provide a good measure of the total group.

A representative sample of the same seventh grade population was given an individual sight-singing test. The data obtained from this individual test was analyzed with the

following conclusions:

1. Twenty-nine percent of the group were skilled in individual sight-singing performance.
2. Thirty-three percent indicated some skill.
3. Thirty-eight percent indicated little or no skill.

A product-moment correlation between group and individual test scores resulted in a .838 coefficient. A more accurate picture of the individual test results was obtained by dividing the group into three parts. A triserial correlation was then possible and resulted in a coefficient of +.954.

CHAPTER VII

SUGGESTIONS FOR FURTHER RESEARCH

CHAPTER VII

SUGGESTIONS FOR FURTHER RESEARCH

1. An item analysis of the group test might prove valuable in determining item difficulty, and aid in revising some of the sub-tests.
2. A revised form of the test should be tried out with a larger number of students from the fourth, fifth and sixth grades. Through proper sampling techniques this group would have a wide representation and suitable norms might be developed.
3. A correlation study between the group test scores and a standardized intelligence test score might be of interest.
4. A correlation study between the group test scores and the Seashore Tests scores would be valuable.
5. Further correlations should be made between group test scores and an established criterion. This criterion may be teacher's estimate of sight-singing ability, or it may be criterion established through a direct approach in the form of an individual sight-singing test. Whatever the method, care should be exercised that the criteria is a good objective measure.

BIBLIOGRAPHY

BIBLIOGRAPHY

Allen, Richard D., Butterfield, Walter H., and Tully, Marguerite, Providence Inventory Test in Music, Yonkers, World Book Company, 1932.

Baker, Earl, and Giddings, T. P., High School Music Teaching, New York, George Banta Publishing Company, 1928.

Beattie, John W., McConathy, Osbourne, and Morgan, Russell V., Music in the Junior High School, Boston, Silver Burdett and Company, 1930.

Book, William F.,"Analysis of the Task of Learning to Read", Journal of Educational Research, 1930, Vol. 21, PP.1-6.

Brooks, Marion B., and Brown, Harry A., Music Education in the Elementary School, New York, American Book Company, 1946.

Brown, A. W.,"The Reliability and Validity of the Seashore Tests of Musical Talent", Journal of Applied Psychology, 1928, Vol. 12, pp. 468-476.

Crossley, Beatrice A., The Construction and Evaluation of a Survey Test in Music for the Primary Grades, unpublished thesis, Boston University, 1946.

Davison, Archibald T., Music Education in America, New York, Harper and Brothers, 1926.

Dean, Charles D., "Predicting Sight-Singing Ability in Teacher Education," Journal of Educational Psychology, November, 1937.

Gehrkens, Karl W., Music in the Grade Schools, Boston, C. C. Birchard Company, 1934.

Gildersleeve, Glenn, Music Achievement Test, Bureau of Publications, Teachers College, Columbia University,1929.

Graham, Percy, "The Function of Rote Singing and Music Reading in the Elementary School," Music Educators National Conference Yearbook, Chicago, 1936.

Hillbrand, E. V., Hillbrand Sight-Singing Test, Yonkers, World Book Company, 1923.

Kerr, Willard A., "Note on Lack of Research in Music," School and Society, 57:476, April, 1943.

Knuth, William E., "Music Research and its Place in the School of Tomorrow," Education, 64:142-145, November, 1943.

Knuth, William E., Knuth Achievement Tests in Music, Minneapolis, Educational Test Bureau., 1936.

Kwalwasser, Jacob, Tests and Measurements in Music, Boston, C. C. Birchard and Company, 1927.

Kwalwasser, Jacob and Dykema, Peter W., K-D Music Tests; Manual of Directions, Boston, Carl Fischer, Inc., 1930.

Kwalwasser, Jacob, Problems in Public School Music, New York, M. Witmark and Sons, 1933.

McConathy, Osbourne, Miessner, W.O., Birge, E. B., and Bray, Mabel E., The Music Hour, Elementary Teacher's Book, Boston, Silver Burdett Company, 1937.

Mosher, R. M., A Study of the Group Method of Measurement of Sight-Singing, Teachers College, Columbia University, Contributions to Education, No. 194, 1925.

Mursell, James L. and Glenn, Mabelle, The Psychology of School Music Teaching, Boston, Silver Burdett Company, 1931.

Mursell, James L., Music in American Schools, Boston, Silver Burdett Company, 1943.

Mursell, James L., Human Values in Music Education, Boston, Silver Burdett Company, 1934.

Norton, Alma, Teaching School Music, New York, C. C. Crawford, 1932.

Peatman, John G., Descriptive and Sampling Statistics, New York, Harper and Brothers, 1947.

Salisbury, Frank, and Smith, Harold B., "Prognosis of Sight-Singing Ability of Normal School Students," Journal of Applied Psychology, October, 1929.

The English Hymnal, London, Oxford University Press, 1946.

Wright, Frances, "The Correlation between Achievement and Capacity in Music," Journal of Educational Research, 1928, Vol. 17, pp. 50-56.

APPENDIX

A. GROUP TEST ADMINISTRATION

B. GROUP TEST SCORING

C. INDIVIDUAL TEST

D. ABSTRACT

Name:_____

Grade:_____ Age:_____

Do you play an instrument?_____

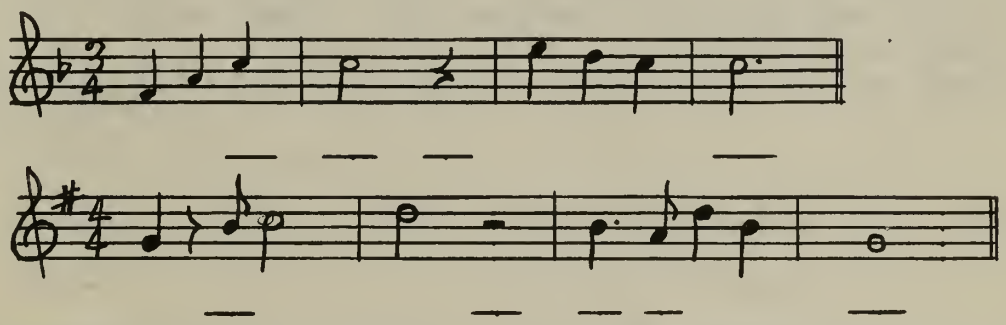
What instrument do you play?_____

How many years have you taken lessons on this instrument?
In class lessons?_____
In private lessons?_____

Part I

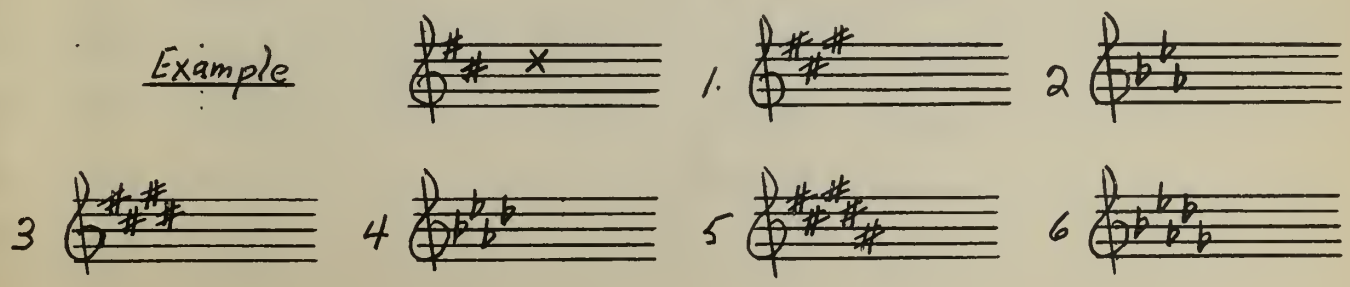
Do not fill in any blank until the teacher says begin. When the teacher says stop you must stop writing whether you have finished or not. If you finish before the teacher says stop wait for instructions. Do not go ahead.

In the following two lines of music you will find blank spaces under some of the notes and rests. You must fill in these blanks with the number of beats you would give the note or rest above the blank. Thus if you decide the note or rest gets one beat you would write 1 in the blank; if the answer is two beats you would write 2.



Part II

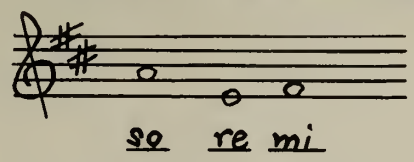
In each of the following you are to place an X on the line or space which is "do" with the given key signature.

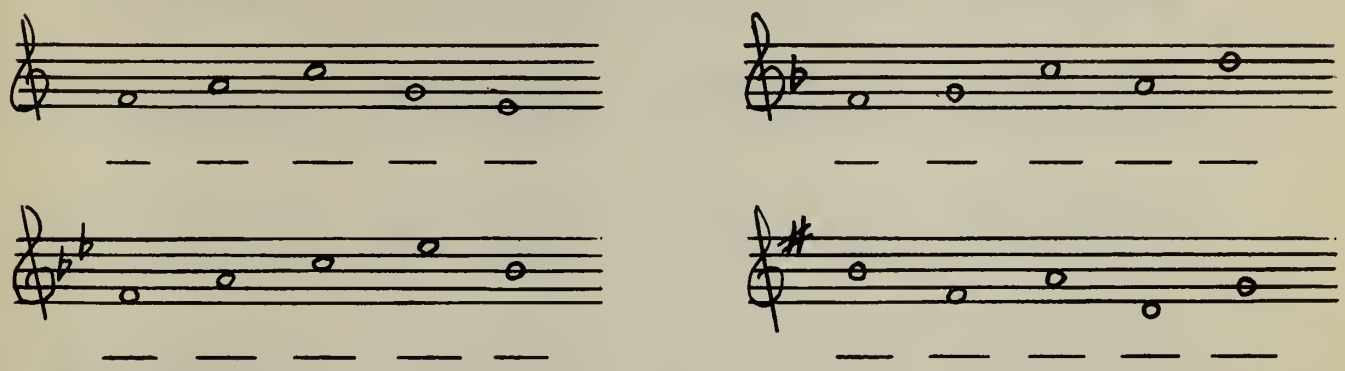


Part III

In the following fill in the blanks below the notes with the proper syllable names. Look at the example.

Example





Part IV

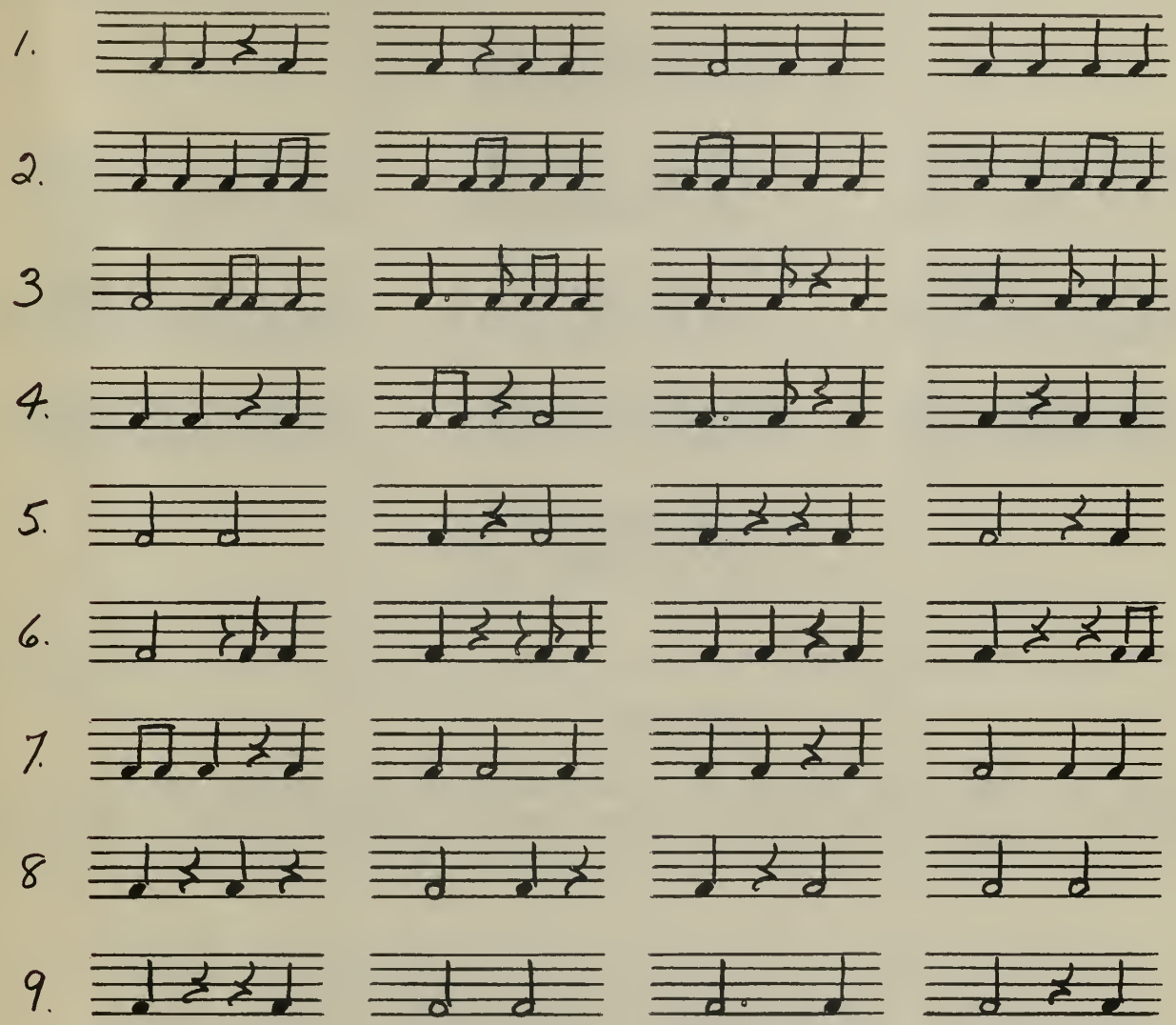
In the first column you see syllables of the scale; opposite each you see a blank space. The teacher will play the pitch of the syllable in the first column and follow it with another tone of the scale, the syllable name of which you must write in the blank space. Look at the example.

Example: so _____

1. do _____
2. do _____
3. re _____
4. so _____
5. do _____
6. la _____
7. mi _____
8. mi _____
9. mi _____
10. ti _____
11. do _____
12. fa _____
13. fa _____
14. re _____
15. so _____
16. do _____
17. mi _____
18. mi _____
19. fa _____
20. re _____

Page V.

On this page you see after each number four groups of notes or notes and rests. Throughout this exercise the quarter note will be equal to one beat. Beginning with group number one (1) the teacher will sing one of the four rhythmic groups; you must pick out the one which the teacher sings and draw a circle around it. Now look at the first group.

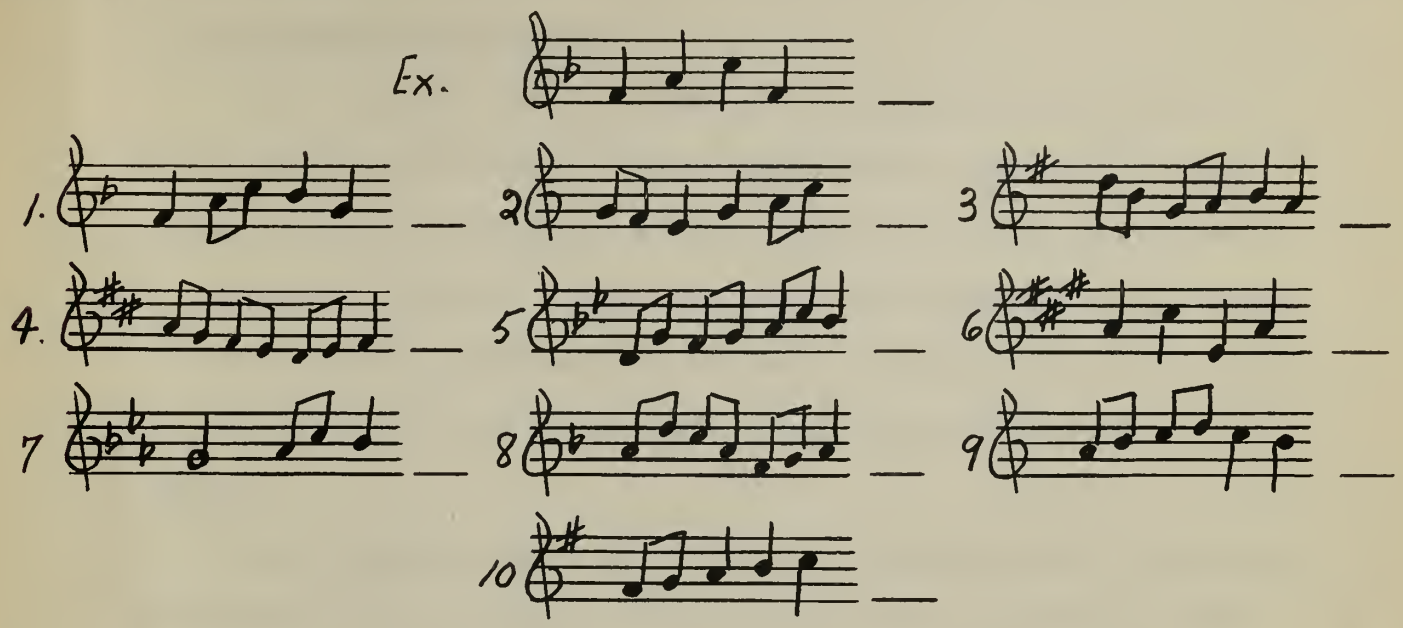


Part VI.

In the following each wave group represents one measure. The beat begins at the top of the wave and continues until the top of the next wave is reached. With a quarter note equal to one beat, you must write the rhythmic exercise which the teacher sings.

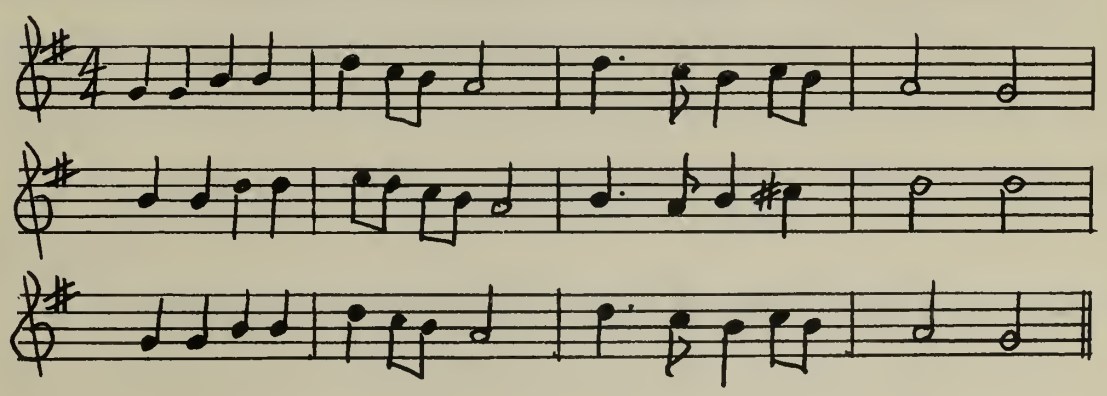
Part VII

The following examples will be played only once but before each you will have a few seconds to look over what is written in order to think of how it should sound. When the example is played you must decide whether it is the same or different from what is written. If you decide it is the same put an S in the blank after the example; if you decide it is different put a D in the blank.



Part VIII

The following hymn tume will be played through only once. While it is being played, if you notice any mistake in pitch or rhythm, you will mark an X through the measure in which the mistake occurred.



APPENDIX

A. GROUP TEST ADMINISTRATION

Preliminary Directions:--

The examiner should take all precautions to insure a good testing situation. Test booklets are given out face down. When the attention of the students has been secured the booklets may be turned over and the students instructed to fill in the information requested at the top of the first page, but not to go ahead until they are instructed.

Part I

The students are instructed to look at Part I and read silently while the teacher reads aloud the directions for Part I. If there are questions from the students, they should be answered if possible by repeating that part of the instructions which was misunderstood. When all understand what to do, they are told to begin. The teacher may proceed to the next part as soon as all have finished, but in no case is more than two minutes allowed for this part.

Part II

Proceed as in Part I allowing no more than two minutes.

Part III

Proceed as in Part I allowing no more than four minutes.

Part IV

The teacher will read the instructions as in previous parts and then say the following: "The tone I play is so

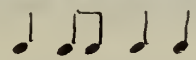
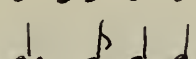
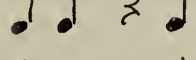
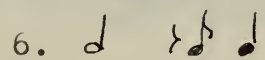
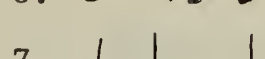
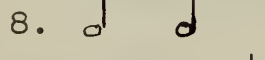
(play so in E); I follow this with another tone (play la in E) the syllable for which you write in the blank opposite; listen to it again (play so-la in E)." "What answer do you have?" "La!" "That's correct." "Throughout this part we will use only one scale which has this pitch for do." (play E fourth space treble staff) "Sing a descending scale from this do down to low do and back up again." "Now look at number one." "I play do (play) and you write the syllable of this tone" (play re) - "again" (play do-re) "Number two" (play do-mi ascending - this and all subsequent intervals are played only once.)

3.	re - so	ascending	10.	ti - la	descending
4.	so - mi	descending	11.	do - so	"
5.	do - ti	"	12.	fa - re	"
6.	la - so	"	13.	fa - mi	"
7.	mi - do	"	14.	re - mi	ascending
8.	mi - so	ascending	15.	so - do	"
9.	mi - fi	"	16.	do - fa	"
			17.	mi - fa	"
			18	mi - la	"
			19.	fa - so	"
			20.	re - do	descending

Part V

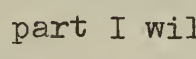
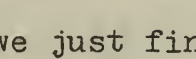
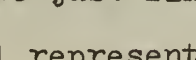
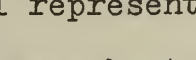
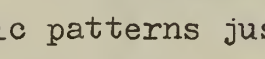
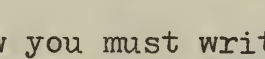
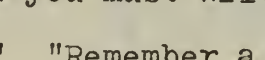
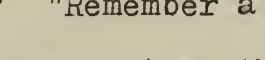
The teacher reads the instructions from the booklet as was done previously and then says: "Before I sing the rhythmic group I will count one measure of four beats to give you the tempo and then immediately sing the example." The teacher then

counts four beats and immediately sings with a neutral syllable this rhythm. This first group is used as an example and does not count in the scoring, so the teacher may check it to be sure the students understand the procedure. Then proceed to sing the following rhythms in like manner.

2. 
3. 
4.
5. 
6. 
7. 
8.
9. 

Part VI

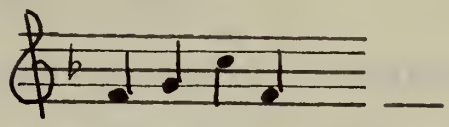
After the teacher reads the instructions aloud, he says: "In this part I will sing rhythmic patterns just as I did in the part we just finished, but now you must write the notes which will represent what I sing." "Remember a quarter note will equal one beat." "Number one: one, two, three, four - ." The teacher may check this first rhythmic pattern on the board to be sure the students understand the procedure which may be unfamiliar. Then proceed to sing the following:

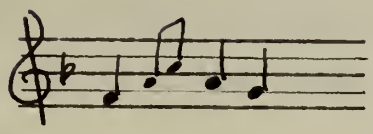
2. 
3. 
4. 
5. 
6. 
7. 
8. 
9. 

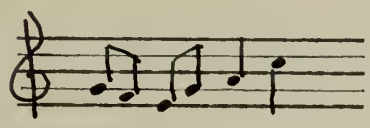
Part VII

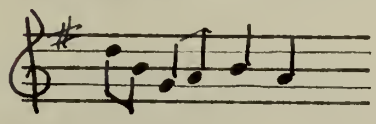
After the teacher reads aloud the instructions, he plays the example and checks the response with the students.

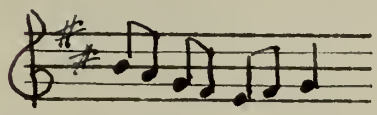
Then proceed to play the following musical figures.

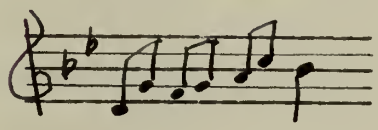
Example 

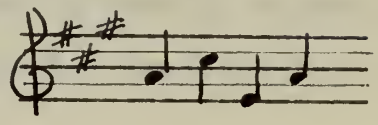
1. 

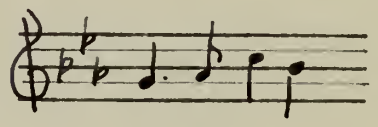
2. 

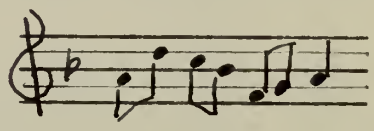
3. 

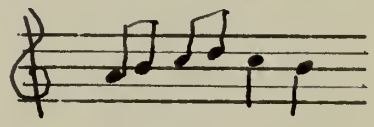
4. 

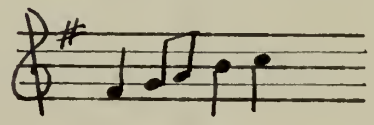
5. 

6. 

7. 

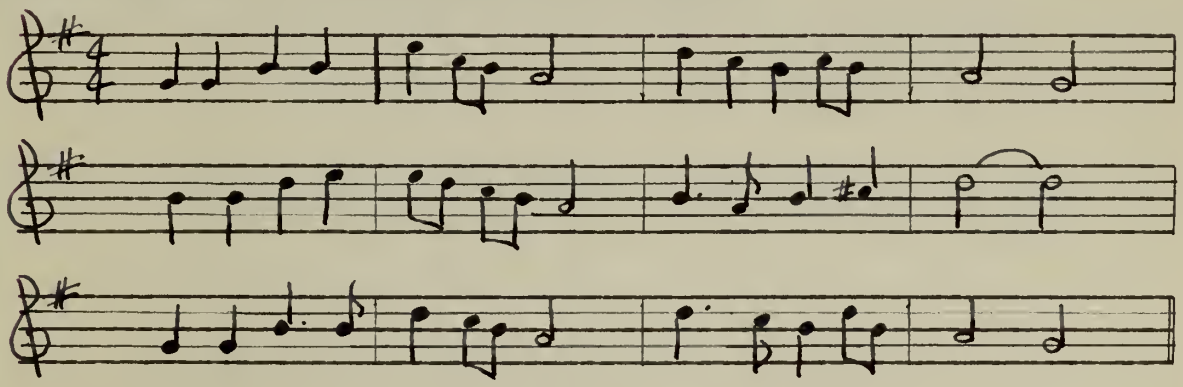
8. 

9. 

10. 

Part VIII

Immediately after reading the instructions aloud, the teacher will count one measure and then play the following hymn tune.



APPENDIX

B. GROUP TEST SCORING

Below in column two is found the number of points credited for each correct response in the sub-test listed in column one. In column three the total possible score for each sub-test is listed.

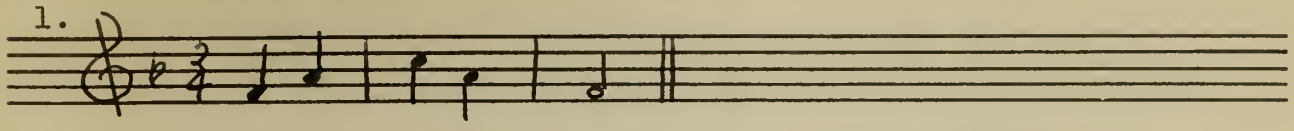
I	II	III
Sub-test I	2	18
Sub-test II	1	6
Sub-test III	1	20
Sub-test IV	2	40
Sub-test V	2	16
Sub-test VI	2	16
Sub-test VII	2	20
Sub-test VIII	4	48
		184

In Sub-test VIII the measure is the unit of measurement. Every measure in the exercise is scored.

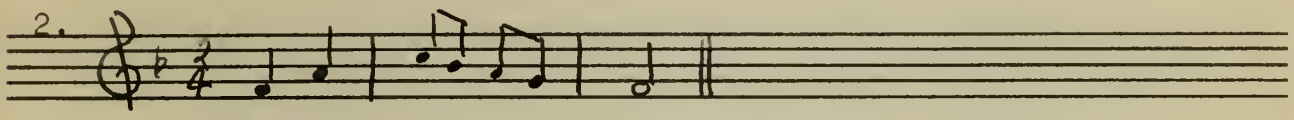
APPENDIX

C. INDIVIDUAL TEST

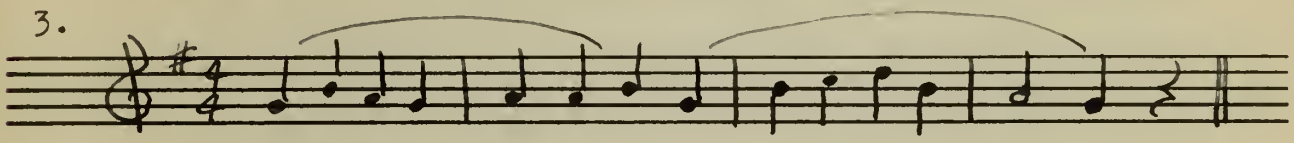
1.



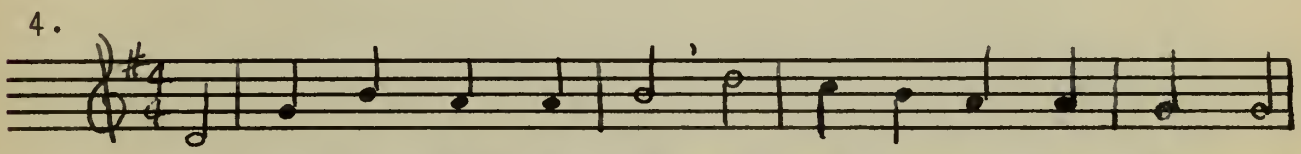
2.



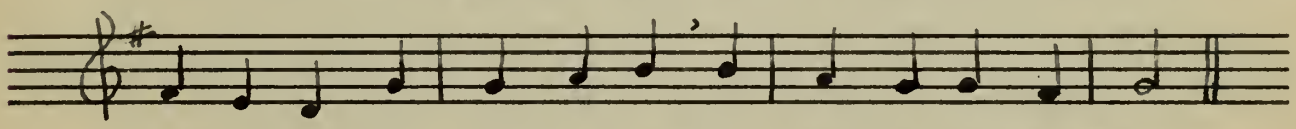
3.



4.

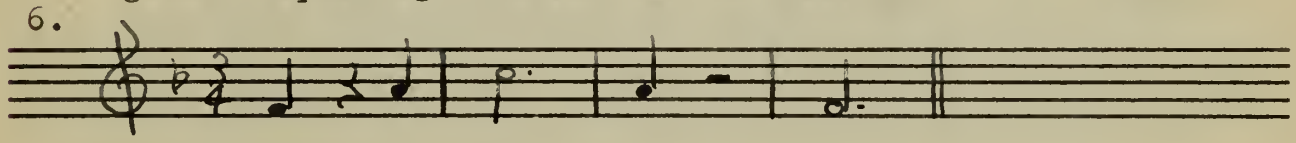


Let sigh-ing cease and woe, God from on high hath heard, the

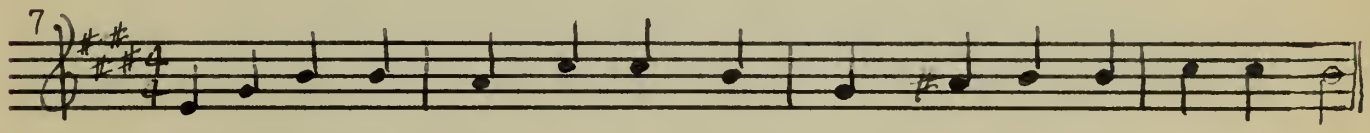


gate is open-ing wide and lo, The long ex-pect-ed Word

6.

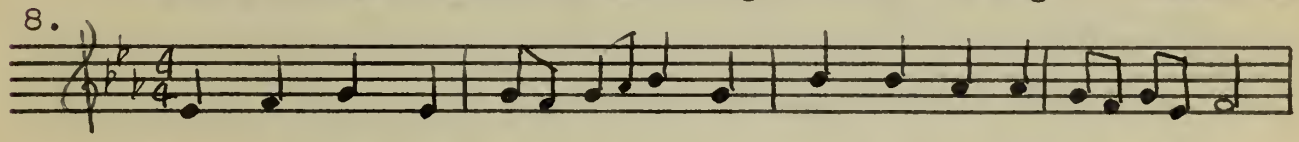


7.

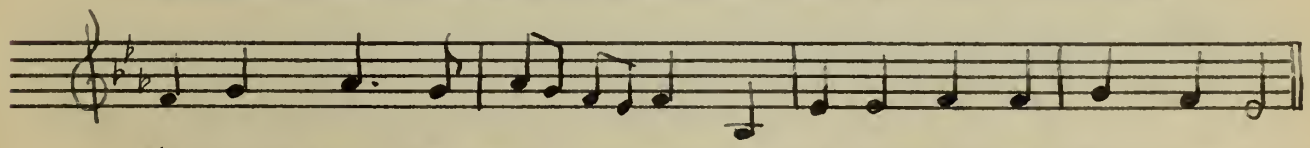


Hark a her-ald voice is call-ing: Christ is nigh it seems to say

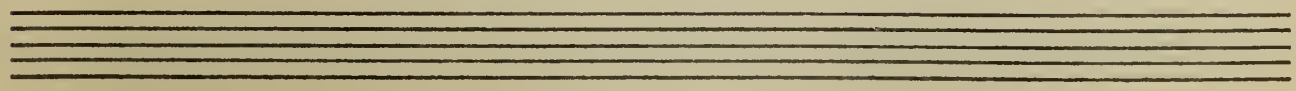
8.



Word sup-reme, be-fore cre-a-tion Born of God e-ter-nal-ly,



Who didst will for our sal-va-tion, To be born on earth to die.



B.F.W.
No. 1F

APPENDIX

C. INDIVIDUAL TEST

In administering the individual test all possible precautions should be taken to insure a good testing situation, so that the students tested will give their best response. A copy of the test is given to the student. When good rapport has been established between student and teacher, the teacher gives the student the first pitch and he proceeds to sing with syllables or with a neutral syllable the first exercise. The procedure for exercise two, three, five and six is similar. In exercises four, seven and eight, the student is encouraged to sing the words, but he may sing these exercises as he did the others with no loss in score.

The teacher keeps score on a copy of the test similar to the copy here. The measure is the scoring unit. If a student makes an error in pitch, a minus sign is placed above the measure; an error in rhythm, a minus sign is placed below the measure. The minus signs are added and subtracted from the total possible score of 76 to get the student's total score. If a student fails completely on the first four exercises there is no need of going further, and he must be scored zero. In such a situation the examining teacher should give much encouragement in order to get the best possible response, but, at the same time, be careful not to offer any more actual help than is given to the better students.

APPENDIX

D. ABSTRACT

Problem:

To construct an objective group test designed to measure several skills which function in the act of sight-singing. Also, to establish a criterion of sight-singing ability against which the group test scores may be correlated to ascertain the efficiency of the group test as a measure of sight-singing.

Scope and limitations of study:

The test consists of elementary sight-singing material at the level of difficulty of simple hymn tunes.

Group test scores were secured from the entire one hundred and eighty seventh grade students in Needham, Massachusetts.

Procedure:

The limits of a simple hymn tune were defined by means of a frequency count of melodic intervals and rhythmic figures of three hundred hymns from the English Hymnal. With this as a basis, a group test of eight parts was constructed and administered. A representative sample group was selected from the tested population. Each student in the sample group was given an individual sight-singing test. The results obtained from this

individual testing form the criterion of sight-singing ability. Correlations were made between the group test scores and the criterion.

Major findings and conclusions:

1. The test proved to be a good measuring instrument for the group.
2. A product-moment linear correlation between the group test scores and the criterion resulted in a +.838 coefficient.
3. A triserial correlation between the group test scores and the criterion resulted in a +.954 coefficient.
4. Twenty-nine percent of the group indicated skill in individual sight-singing.
5. Thirty-three percent of the group indicated some skill.
6. Thirty-eight percent indicated little or no skill in sight-singing.

BOSTON UNIVERSITY
1 1719 02561 3185